Soluautomaattien approksimaatioista

Ilmari Karonen

4. maaliskuuta 2009

LuK-tutkielma

Johdanto

Tässä tutkielmassa tarkastelen soluautomaatteja ja niiden yleisesti käytettyjä keskivertokenttä- ("mean field") ja pariapproksimaatioita. Määrittelen soluautomaatin käsitteen yleisellä tasolla ja kuvailen kolme eri tyypillisesti käytettyä soluautomaattityyppiä, joista kahdesta esimerkkeinä mainitsen John Conwayn tunnetun "Game of Life" -soluautomaatin sekä epidemiologian alalla yleisesti käytetyn hilakontaktiprosessin. Tutkielman toisessa osassa esittelen eri tapoja, joilla soluautomaattien käyttäytymistä voidaan analyyttisesti approksimoida ja karakterisoida, mukaanlukien niin kutsuttu soluautomaatin aktiviteetti (λ) sekä yllämainitut keskivertokenttä- ja pariapproksimaatiot. Kahta jälkimmäistä sovellan hilakontaktiprosessiin, ja vertailen näin saatuja tuloksia.

Sisältö

1 Soluautomaatit

Soluautomaatteja käytetään erilaisten spatiaalisten prosessien mallintamiseen monilla aloilla fysiikasta biomatematiikkaan, minkä lisäksi ne itsessään ovat olleet kiinnostuksen kohteena mm. laskentateoreettisten ominaisuuksiensa vuoksi. Eri aloilla soluautomaattien käsite on tapana määritellä usein hieman eri tavoilla. Sen vuoksi lienee syytä määritellä aluksi, mitä soluautomaateilla tässä tutkielmassa tarkoitan.

1.1 Määritelmä

Soluautomaatit, kuten nimistä voi päätellä, koostuvat joukosta *soluja*, jotka muodostavat niin kutsutun *hilan* L. Kullakin hilan solulla $x \in L$ on määrätty joukko *naapureita*, $N_x \subset L$, joiden kanssa se on vuorovaikutuksessa. (On määritelmäkysymys, katsotaanko kukin solu itsensä naapuriksi vai ei. Koska molemmilla määritelmillä on puolensa, käytän tässä merkintää N_x solun x naapurustosta ilman solua x itseään ja merkintää $\bar{N}_x = N_x \cup \{x\}$ naapurustosta x mukaanlukien.)

Kullakin ajanhetkellä kukin solu on yhdessä äärellisestä joukosta tiloja Ω.[1] Käytän merkintää $\sigma_x(t) \in \Omega$ vittaamaan solun $x \in L$ tilaan ajanhetkellä t, jotka yhdessä muodostavat automaatin kokonaistilan $\sigma(t)$. Solujen tila kehittyy ajan myötä kullekin soluautomaatille ominaisen päivityssäännön mukaan. Päivityssääntöjä on monenlaisia, mutta yhteistä kaikille soluautomaateille on se, että kukin solun tilan muutokset riippuvat ainoastaan solun itsensä ja sen naapurisolujen tilasta kyseisellä ajanhetkellä – ei kauempana sijaitsevista soluista (*paikallisuus*) eikä solujen mahdollisista aiemmista tiloista (*muistittomuus*).

Solujen tilan kehityksen määräävät päivityssäännöt voidaan jakaa useaan luokkaan. Perinteisimmissä soluautomaateissa solujen päivitys tapahtuu *samanaikaisesti* ja *deterministisesti* – t.s. ilman satunnaisuutta – niin, että kunkin solun x tila ajanhetkellä $t+1$ määräytyy yksikäsitteisesti solun itsensä ja sen naapurisolujen tilan ajanhetkellä t mukaan:

$$\sigma_x(t+1) = r_x(\sigma_{\bar{N}_x}(t)). \tag{1.1}$$

Monissa käytännön sovelluksissa on hyödyksi sallia solujen päivityksen olevan myös jossain määrin satttumanvaraista. Tällaisissa *stokastisissa* soluautomaateissa kunkin solun ja sen ympäristön tila hetkellä t määrää *todennäköisyyden*, jolla solu tulee olemaan kussakin mahdollisessa tilassa hetkellä

[1] Sopivin muutoksin voisi tilajoukon sallia olevan myös ääretön, esim. $\mathbb{Z}$ tai $\mathbb{R}$, mutta tämä monimutkaistaisi myöhemmin esiteltäviä approksimaatiomenetelmiä vaatien niissä esiintyvien äärellisten summien korvaamista äärettömillä summilla tai integraaleilla.

$t+1$. Muodollisesti, kaikille mahdollisille tiloille $a \in \Omega$,

$$\Pr[\sigma_x(t+1) = a] = r_x(a|\sigma_{\bar{N}_x}(t)). \tag{1.2}$$

Ylläoleva sääntö, sovellettuna kuhunkin hilan soluun, määrittää automaatin kokonaistilan stokastisen kehitysradan. Ottamalla yhtälöstä 1.2 summa kaikkien automaatin mahdollisten kokonaistiloihin $\tau \in \Omega_L$ yli saadaan kuitenkin yhtälö, joka kuvaa automaatin eri tilojen *todennäköisyyksien* determinististä kehitystä:

$$\Pr[\sigma(t+1) = \rho] = \sum_{\tau \in \Omega_L} \Pr[\sigma(t) = \tau] \prod_{x \in L} r_x(\rho_x|\tau_{\bar{N}_x}), \tag{1.3}$$

missä ρ_x on solun x tila kokonaistilassa ρ ja $\tau_{\bar{N}_x}$ on solun x ja sen naapuruston tila kokonaistilassa τ.

Molemmissa edellämainituissa automaattityypeissä kaikkien solujen päivitys tapahtuu samanaikaisesti. Toinen mahdollisuus on sallia kunkin solun vaihtaa tilaansa millä ajanhetkellä tahansa päivityssäännön määrätessä tapahtumatiheyden, jolla solu senhetkisestä naapurustostaan riippuen siirtyy kustakin tilasta toiseen. Jos muutostapahtumat oletetaan toisistaan riippumattomiksi, voidaan päivityssääntö esittää joukkona differentiaaliyhtälöitä niin, että, merkiten $p(\tau, t) = \Pr[\sigma(t) = \tau]$ ja $p_x(a, t) = \Pr[\sigma_x(t) = a]$, kaikille tiloilla $a \in \Omega$ pätee

$$\frac{d}{dt} p_x(a) = \sum_{\tau \in \Omega_L} p(\tau, t) \sum_{b \in \Omega} (f_x(b \to a|\tau_{N_x}) - f_x(a \to b|\tau_{N_x})), \tag{1.4}$$

missä

$$f_x(a \to b|\tau_{N_x}) = p_x(a, t)\ r_x(a \to b|\tau_{N_x}) \tag{1.5}$$

ja $r_x(a \to b|\tau_{N_x})$ on tiheys, jolla solu x siirtyy tilasta a tilaan b sen naapuruston ollessa tilassa τ_{N_x}.[2] Kuten edellä, siirtymätiheysfunktio $r_x(a \to b|\tau_{N_x})$ määrittelee yhtälöiden 1.4 ja 1.5 kautta yksikäsitteisesti automaatin eri kokonaistilojen todennäköisyysjakauman kehityksen ajan myötä.

1.2 Esimerkkejä

Jottei yllä annettu määritelmä jäisi turhan abstraktiksi, annan vielä muutaman esimerkin eri tyyppisistä soluautomaateista.

[2] Olisi mahdollista tarkastella myös malleja, joissa useampi vierekkäinen solu voi vaihtaa tilaa samanaikaisesti, esim. niin että kaksi vierekkäistä solua vaihtaa tilaa keskenään, mutta tällöin yhtälöissä 1.4 ja 1.5 täytyisi käsitellä useamman solun ryhmiä.

Esimerkki 1.1 (Conwayn "Game of Life")**.** Ehkä tunnetuin esimerkki "perinteisestä" deterministinisestä soluautomaatista on John Conwayn vuonna 1970 kehittämä niin kutsuttu "Game of Life" ([Gar70]). Tämän "elämänpeli" koostuu (periaatteessa) äärettömästä kaksiulotteisesti ruudukosta soluja ($L = \mathbb{Z}^2$), joista kukin on yhdessä kahdesta tilasta, jotka on perinteisesti nimetty $\Omega = \{elossa, kuollut\}$.[3] Kunkin solun naapureita ovat sen neljä vaaka- ja pystysuoraan ja neljä vinottain viereistä solua. Päivityssääntö on, että kukin solu on *elossa* hetkellä $t + 1$ jos ja vain jos hetkellä t joko

(a) se itse on *kuollut* ja tasan kolme sen kahdeksasta naapurista on *elossa*, tai

(b) se itse on *elossa* ja kaksi tai kolme sen kahdeksasta naapurista on *elossa*.

Tämä yksinkertainen sääntö johtaa yllättävän monimutkaiseen käyttäytymiseen. Useimmista rajallisen määrän *elossa* olevia soluja sisältävistä lähtötilanteista päivityssäännön toistuva soveltaminen johtaa, mahdollisesti hyvinkin pitkän ajan kuluttua, lopulta joukkoon erillisiä stabiileja tai oskilloivia soluryhmiä. On kuitenkin mahdollista löytää myös lähtöasetelmia jotka kasvavat äärettömästi, esim. muodostamalla päättymättömän jonon n.k. *glidereita*, viiden elävän solun ryhmiä jotka neljän päivityksen jälkeen palaavat alkutilaansa yhden solunmitan vinottain siirtyneinä.

Törmätessään toisiinsa tai muihin liikkuviin tai paikallaan pysyviin soluryhmiin tällaiset *gliderit* voivat saada aikaan monia erilaisia tapahtumaketjuja, joista jotkin johtavat esim. toisen *gliderin* häviämiseen tai useamman uuden *gliderin* muodostumiseen. Käyttämällä tällaisia törmäyksiä loogisten operaattoreiden mallintamiseen on voitu osoittaa, että "Game of Life" on *laskennallisesti universaali* – sopivasta alkuasetelmasta tämä soluautomaatti kykenee teoriassa ratkaisemaan minkä tahansa laskentatehtävän, joka on Turingin koneen laskettavissa ([Ada02]).

Esimerkki 1.2 (Hilakontaktiprosessi)**.** Kontaktiprosesseja käytetään yleisesti epidemiologian alalla mallintamaan tartuntatautien leviämista populaatiossa. Yksinkertaisimmassa, niin kutsutussa "SIS"-kontaktiprosessissa populaatio koostuu suuresta joukosta yksilöitä, jotka jakaantuvat kahteen luokkaan, tartunnalle alttiisiin (S = *susceptible*) ja tartunnan saaneisiin (I = *infected*), ja ovat kaikki muuten identtisiä. Kukin yksilö kohtaa c muuta yksilöä aikayksikköä kohden; kun tartunnalle altis yksilö kohtaa tartunnan saaneen,

[3] Conway simuloi aluksi kehittämäänsä peliä käsin *go*-laudalla, käyttäen pelin mustia ja valkoisia nappuloita esittämään solujen eri tiloja.

se saa tältä tartunnan todennäköisyydellä p. Tartunnan saaneet yksilöt paranevat tiheydellä r, ja ovat parannuttuaan jälleen alttiita uudelle tartunnalle.

Jos populaation oletetaan olevan hyvin sekoittunut, niin että kaikilla yksilöpareilla on yhtä suuri todennäköisyys kohdata toisensa, saadaan populaation koon lähestyessä ääretöntä keskimääräisen tartuntatiheyden i kehitykselle seuraava deterministinen kaava:

$$\frac{\mathrm{d}i}{\mathrm{d}t} = cp(1-i)i - ri \tag{1.6}$$

Tällä differentiaaliyhtälöllä on triviaali kiintopiste arvolla $i = 0$. Jos keskimääräinen tarttumistiheys cp ei ylitä keskimääräistä parantumistiheyttä r, on tämä triviaali kiintopiste yhtälön ainoa, ja koska $\mathrm{d}i/\mathrm{d}t < 0$ kaikille $i > 0$, hakeutuu i tähän kiintopisteeseen kaikista lähtöarvoista $0 \le i_0 \le 1$. Jos $cp > r$, on triviaali kiintopiste sen sijaan epästabiili, ja i lähestyy sen sijaan arvoa $\tilde{\imath} = 1 - r/cp$.

Saadaksemme malliin mukaan konkreettisen spatiaalisen elementin, voimme olettaa yksilöiden sijaitsevan jollain sopivasti valitulla hilalla L, niin että kullakin yksilöllä on n naapuria, ja kohtaavan ainostaan naapureitaan tällä hilalla. Näillä oletuksilla SIS-mallista muodostuu jatkuva-aikainen soluautomaatti, jonka solujen tilan kehityksen määrää sääntö

$$\frac{d}{dt}\Pr[\sigma_x = \mathrm{I}] = cp \sum_{\tau\in\Omega_L} \Pr[\sigma(t) = \tau]\delta_{\mathrm{S}}(\tau_x)\frac{n_{\mathrm{I}}(\tau_{N_x})}{n} - r\Pr[\sigma_x(t) = \mathrm{I}], \tag{1.7}$$

missä $\delta_{\mathrm{S}}(\tau_x)$ on 1 jos $\tau_x = \mathrm{S}$ ja muuten 0, ja $n_{\mathrm{I}}(\tau_{N_x})$ on I-tilassa olevien solujen lukumäärä solun x naapurustossa N_x. Tämä sääntö voidaan ilmaista yhtälöiden 1.4 ja 1.5 muodossa asettamalla

$$r_x(\mathrm{S} \to \mathrm{I}|\tau_{N_x}) = cp\frac{n_{\mathrm{I}}(\tau_{N_x})}{n}, \text{ ja} \tag{1.8}$$

$$r_x(\mathrm{I} \to \mathrm{S}|\tau_{N_x}) = r. \tag{1.9}$$

Jos hilan oletetaan koostuvan äärettömästä määrästä soluja, kehittyy tilassa I olevien solujen keskimääräinen tiheys i jälleen deterministisesti. Yleispiirteiltään kontaktiprosessin kehitys hilalla osoittautuu monin tavoin hyvin sekoittuneen populaation tapausta muistuttavaksi: kriittisen parametrin r/cp arvon ollessa korkea tartuntatiheys i putoaa nollaan, kun taas parametrin alemmilla arvoilla se lähestyy epätriviaalia kiintopistettä $0 < \tilde{\imath} \le 1$. Solujen välisten kohtaamismahdollisuuksien paikallisuus aiheuttaa kuitenkin tartuntojen kasaantumista, mikä laskee S–I -kohtaamisten todennäköisyyttä ja täten hidastaa tartunnan leviämistä. Tämä laskee sekä keskimääräistä tartuntatiheyttä että tartunnan häviämiseen vaadittavaa parametrin r/cp arvoa hyvin sekoittuneeseen populaatioon verrattuna. Lopullinen tartuntatiheys hilakontaktiprosessissa riippuu sekä kunkin solun naapurien lukumäärästä että hilan geometriasta.

2 Soluautomaattien approksimointi

Soluautomaatit ovat hyödyllisiä malleja monenlaisille spatiaalisille ilmiöille, mutta niiden ongelmana on usein juuri sama rikkaus, joka tekee niistä kiinostavia: monien soluautomaattien kehitystä on vaikeaa ellei mahdotonta ennustaa muuten kuin kokeilemalla, ja tällöinkin luotettavien tulosten saaminen vaatii usein huomattavan pitkiä ja laajoja simulaatioita useasti toistettuina. Tämän vuoksi on kehitetty erilaisia menetelmiä, joilla soluautomaattien käyttäytymistä voidaan pyrkiä ainakin karkeasti approksimoimaan analyyttisin menetelmin.

2.1 Aktiviteetti

Kaikkein yksinkertaisin tapa arvioida soluautomaatin käyttäytymistä on sen niin kutsuttu *aktiviteetti*, jolle kirjallisuudessa usein käytetään symbolia λ (ks. esim. [Wol02]). Tämä on yksinkertaisesti lukuarvo, joka kertoo annetun solun todennäköisyys päätyä "aktiiviseen" tilaan summattuna kaikkien mahdollisten solun ja sen naapuruston tilojen yli. Sitä, mitkä tilat katsotaan "aktiivisiksi", ei ole tarkemmin määritelty, eikä luontevaa määritelmää kaikille soluautomaateille voitane antaakaan. Mikäli automaatilla kuitenkin on jokin "neutraali" tila – kuten esim. *kuollut* tila "Game of Lifessä" tai S-tila hilakontaktiprosessissa – jossa olevat solut pysyvät samassa tilassa, ellei ainakin joku niiden naapureista ole jossain muussa tilassa, on usein luontevaa kutsua tätä tilaa "passiviiseksi" ja muita mahdollisia tiloja "aktiivisiksi". Aktiviteetin käsite voidaan myös yleistää laskemalla yllä kuvattu todennäköisyys jokaiselle tilalle erikseen.

Aktiviteetin voidaan tulkita kertovan, miltä hila keskimäärin näyttäisi *yhden* päivityksen jälkeen, jos solujen tilat olisivat alunperin tasaisesti ja toisistaan riippumattomasti jakaantuneita. Vaikka tämä karkea luonnehdinta selvästikin kertoo hyvin vähän automaatin todellisesta käyttäytymisestä, on sitä silti käytetty (lähinnä perinteisiä deterministisiä soluautomaatteja tutkivien laskentateoreetikkojen parissa) jopa jonkinasteisella menestyksellä ennustamaan automaattien käyttäytymisen monimutkaisuutta. Aktiviteetin etuna tässä on sen yksinkertaisuus: se voidaan laskea hyvin suoraviivaisesti suljetussa muodossa automaatin sääntötaulukosta.

2.2 Keskivertokenttäapproksimaatio

Yksinkertaisin approksimaatiomenetelmä, jonka voidaan todella katsoa kertovan jotain automaatin käyttäytymisestä ajan myötä, on niin kutsuttu keskivertokenttä- eli *mean field* -approksimaatio.[4] Tämä approksimaatio perustuu oletukseen, että hilan solujen tilat ovat toisistaan riippumattomasti jakaantuneet saman keskimääräisen todennäköisyysjakauman mukaan. Tällä oletuksella kunkin soluryhmän X todennäköisyys olla tilassa $A \in \Omega_X$ voidaan kirjoittaa suljetussa muodossa

$$\Pr[\sigma_X = A] = \prod_{y \in X} p_{\mathrm{mf}}(A_y), \tag{2.1}$$

missä A_y on solun y tila ryhmän X kokonaistilassa A ja $p_{\mathrm{mf}}(a)$ on tilan $a \in \Omega$ keskivertotodennäköisyys.

Soveltamalla yhtälöä 2.1 yhtälöön 1.3 saadaan (mahdollisesti stokastisen) synkronisen soluautomaatin keskivertotodennäköisyyksien kehitykselle yhtälö

$$p_{\mathrm{mf}}(b; t+1) = \sum_{A \in \Omega_{\bar{N}}} r(b|A) \prod_{y \in \bar{N}} p_{\mathrm{mf}}(A_y; t), \tag{2.2}$$

missä $\bar{N}$ on vapaasti valitun solun naapurusto (jotka kaikki oletetaan keskenään isomorfisiksi) solu itse mukaanlukien ja $\Omega_{\bar{N}}$ sen mahdollisten kokonaistilojen joukko. Vastaavasti yhtälöistä 1.4 ja 1.5 saadaan samoin oletuksin jatkuva-aikaisen soluautomaatin keskivertotodennäköisyyksien kehitykselle differentiaaliyhtälö

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{\mathrm{mf}}(b) = \sum_{a \in \Omega, C \in \Omega_N} \phi(a \rightleftharpoons b|C) \prod_{y \in N} p_{\mathrm{mf}}(C_y), \tag{2.3}$$

missä N on vapaasti valitun solun naapurusto, tällä kertaa ilman solua itseään, Ω_N sen mahdollisten tilojen joukko, ja

$$\phi(a \rightleftharpoons b|C) = p_{\mathrm{mf}}(a) r(a \to b|C) - p_{\mathrm{mf}}(b) r(b \to a|C) \tag{2.4}$$

on keskimääräinen siirtymätiheys tilojen a ja b välillä ympäristössä C.

Keskivertokenttäapproksimaation soveltaminen kappaleessa 1.2 kuvattuun hilakontaktiprosessiin antaa tulokseksi täsmälleen saman mallin kuin kontaktiprosessia hyvin sekoittuneessa populaatiossa kuvaava yhtälö 1.6. Tämä ei itse asiassa ole mikään yllätys, sillä keskivertokenttämallin voidaan tulkita

[4] Termi “mean field” on peräisin tilastollisen fysiikan alalta, jossa vastaavantyyppisiä approksimaatioita käytetään keskenään vuorovaikuttavien kappaleiden käyttäytymisen mallintamiseen olettamalla, että kuhunkin kappaleeseen vaikuttavia voimia voidaan approksimoida kaikkien kappaleiden mahdollisten keskinäisten sijaintien yli summatulla keskimääräisellä voimakentällä. Suomenkielinen termi “keskivertokenttäapproksimaatio” on kirjoittajan oma käännösyritys.

kuvaavan tilannetta, jossa hilan solut vaihtavat jatkuvasti paikkoja keskenään satunnaisesti ja nopeaan tahtiin. Täten keskivertokenttäapproksimaatiota voidaan käyttää johtamaan mille tahansa annetulle soluautomaatille "hyvin sekoitettu" muunnelma, jota voidaan pitää automaatin kanonisena epäspatiaalisena vastineena.

2.3 Pariapproksimaatio

Edellä käsitelty keskivertokenttämalli jättää lähes täysin huomiotta soluautomaatin hilarakenteen. Haluttaessa tutkia edes jollain tasolla hilan vaikutusta automaatin käyttäytymiseen voidaan siirtyä ns. pariapproksimaatioon (ks. esim. [vB00]), joka muodostetaan muuten samalla tavoin, mutta jossa yksittäisten solujen sijasta tarkastellaan naapurisolujen muodostamia pareja ja niiden yhteisten tilojen todennäköisyyksiä.

Sen sijaan, että solujen tilatodennäköisyydet oletettaisiin toisistaan riippumattomiksi (mikä johtaisi keskivertokenttämalliin), voidaan tässä tapauksessa approksimaatio-oletuksena käyttää *ehdollista riippumattomuutta*, jonka mukaan kunkin solun tilatodennäköisyydet riippuvat ainoastaan sen naapurisolujen tiloista. Tästä oletuksesta ei suoraan voida johtaa yhtälön 2.1 kaltaista lauseketta vapaavalintaisten soluryhmien kokonaistilatodennäköisyyksille kuin joissain erikoistapauksissa,[5] mutta olettamalla naapurisolujen naapurustot soluja itseään lukuunottamatta erillisiksi, ja olettamalla lisäksi, että kunkin solun naapurien tilat ovat toisistaan (ehdollisesti) riippumattomia, voidaan yksittäisen soluparin xy naapuruston $\hat{N}_{xy} = xy \cup (N_x \setminus y) \cup (N_y \setminus x)$ kokonaistilojen todennäköisyyksille johtaa kaava

$$\Pr[\sigma_{\hat{N}_{xy}} = A] = p(A_{xy}) \prod_{r \in N_x \setminus y} p(A_r | A_x) \prod_{s \in N_y \setminus x} p(A_s | A_y), \tag{2.5}$$

missä $p(A_{xy})$ on soluparin xy todennäköisyys olla tilassa A_{xy}, ja $p(A_r|A_x) = p(A_{rx})/p(A_x)$ on solun r ehdollinen todennäköisyys olla tilassa A_r sen naapurisolun x ollessa tilassa A_x.

Yhtälöä 2.5 voidaan kuten yllä solveltaa yhtälöön 1.3. Täten saadaan (mahdollisesti stokastisen) synkronisen soluautomaatin paritodennäköisyyk-

[5]Niin sanotuille Bethe-hiloille, joihin kuuluvat mm. yksiulotteiset hilat ja satunnaisverkot, tällainen lauseke voidaan antaa yksinkertaisella rekursiokaavalla. Perusvaatimuksena tälle on, että äärelliset soluryhmät eivät saa sisältää pareista muodostuvia silmukoita, jolloin yhtälön 2.5 oletukset toteutuvat kaikkialla. Yleisessä tapauksessa vastaava kaava voidaan periaatteessa muodostaa variaatiolaskennan kautta, maksimoimalla kokonaistilatodennäköisyyksien Shannon-informaatio paritodennäköisyyksien antamien reunaehtojen puitteissa, mutta tämä johtaa äkkiä hyvin monimutkaisiin ja analyyttisesti ratkeamattomiin lausekkeisiin.

sien kehitykselle yhtälö

$$(2.6) \qquad p(bc;t+1) = \sum_{A\in\Omega_{\bar{N}_{xy}}} r(b|A_{\hat{N}_x})r(c|A_{\hat{N}_y})p_{\text{ext}}(A;t),$$

missä xy on vapaavalintainen solupari, $\bar{N}_{xy}$ sen naapurusto (jotka kaikki oletetaan jälleen keskenään isomorfisiksi) pari itse mukaanlukien, $\Omega_{\bar{N}_{xy}}$ sen mahdollisten kokonaistilojen joukko ja $p_{\text{ext}}(A;t) = \Pr[\sigma_{\hat{N}_{xy}}(t) = A]$ on yhtälön 2.5 mukaan laskettu tilan A todennäköisyys hetkellä t. Vastaavasti yhtälöistä 1.4 ja 1.5 saadaan yhtälöä 2.5 soveltamalla jatkuva-aikaisen soluautomaatin paritodennäköisyyksien kehitykselle differentiaaliyhtälö

$$(2.7) \qquad \frac{\mathrm{d}}{\mathrm{d}t}p(bc) = \sum_{a\in\Omega, D\in\Omega_{N_{xy}}} \theta(a \rightleftharpoons b|c,D) + \theta(a \rightleftharpoons c|b,D)$$

missä

$$(2.8) \qquad \theta(a \rightleftharpoons b|c,D) = p_{\text{ext}}(acD)r(a\to b|cD) - p_{\text{ext}}(bcD)r(b\to a|cD)$$

on parin xy yhden solun keskimääräinen siirtymätiheys tilojen a ja b välillä parin toisen solun ollessa tilassa c ja parin naapuruston ollessa tilassa D kerottuna yhtälön 2.5 mukaan lasketuilla alkutilojen todennäköisyyksillä.

Kappaleen 1.2 hilakontaktiprosessille pariapproksimaatio antaa yhtälöstä 1.6 poikkeavan tuloksen

$$(2.9) \qquad \frac{\mathrm{d}i}{\mathrm{d}t} = cp(i-j) - ri$$

$$(2.10) \qquad \frac{\mathrm{d}j}{\mathrm{d}t} = 2cp\frac{i-j+(n-2)i^2-(2n-3)ij+(n-1)j^2}{(1-i)n} - 2rj,$$

missä i on satunnaisesti valitun solun todennäköisyys olla tartunnan saanut, kuten yhtälössä 1.6, j todennäköisyys, että satunnaisesti valitun naapurisoluparin molemmat solut ovat tartunnan saaneita, ja n kunkin solun naapureiden lukumäärä. Yhtälöiden ratkaisu antaa tulokseksi, että tartunta on elinkelpoinen jos $r/cp < (1-n)/n$ (vertaa hyvin sekoittuneen populaation vastaavaan ehtoon $r/cp < 1$), ja että tällöin keskimääräinen tartunnan saaneiden tiheys lähestyy ajan myötä arvoa

$$(2.11) \qquad \tilde{\imath} = \frac{n-1-nr/cp}{n-1-r/cp} = \frac{cp(n-1)-nr}{cp(n-1)-r}.$$

Naapurien lukumäärän n kasvaessa tämä lähestyy hyvin sekoittuneen populaation tapauksessa saatua vastaavaa arvoa $1 - r/cp$.

Pariapproksimaatiolle on jossain määrin hankalampi antaa samanlaista konkreettista tulkintaa, kuin mikä keskivertokenttäapproksimaatiolle voitiin antaa solujen nopean sekoittumisen kautta. Jonkinlainen tulkinta voidaan ainakin silmukattomien hilojen tapauksissa antaa olettamalla, että hila voi "ristetä" kahden samantilaisen soluparin xy ja zw kohdalta niin, että aiemmat parit katkeavat ja solut xw ja zy muodostavat uudet parit, ja että tällaisia risteymiä tapahtuu jatkuvasti ja nopeaan tahtiin. Tämä tulkinta on silti selvästi työstetympi kuin vastaava keskivertokenttäapproksimaatiolle annettu, eikä se siltikään sovellu esim. kaksi- tai useampiulotteisille hiloille, joissa soluparit voivat muodostaa silmukoita. Toisaalta jo yhtälöä 2.5 johdettaessa tehtiin oletuksia, jotka eivät näillä hiloilla toteudu; yksi tapa tulkita asiaa onkin, ettei yhtälön 2.5 antama klassinen pariapproksimaatio kuvaa näitä hiloja ensinkään, vaan se oikeastaan approksimoi automaatin käyttäytymistä vastaavanasteisella satunnaisverkolla.

Merkinnät ja määritelmät

$\mathbb{Z}$	Kokonaislukujen joukko: $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{R}$	Reaalilukujen joukko.
L	Soluautomaatin hila.
$x, y, z \in L$	Yksittäisiä soluja.
$X, Y, Z \subset L$	Soluryhmiä.
$xy := \{x, y\}$	Solupari.
$XY := X \cup Y$	Kahden (yleensä erillisen) soluryhmän yhdiste.
N_x	Solun x naapurisolujen joukko eli naapurusto, poislukien solu x itse.
$\bar{N}_x := N_x \cup \{x\}$	Solun x naapurisolujen joukko eli naapurusto, mukaanlukien solu x itse.
$N_X := \bar{N}_x \setminus X$	Soluryhmän X naapurusto, poislukien soluryhmä X itse.
$\bar{N}_X := \bigcup_{x \in X} \bar{N}_x$	Soluryhmän X naapurusto, mukaanlukien soluryhmä X itse.
$\sigma = \sigma(t) = \sigma_L(t)$	Soluautomaatin kokonaistila ajanhetkellä t.
$\sigma_x = \sigma_x(t)$	Solun $x \in L$ tila ajanhetkellä t.
$\sigma_X = \sigma_X(t)$	Soluryhmän $X \subset L$ tila ajanhetkellä t.
Ω	Kunkin yksittäisen solun mahdollisten tilojen joukko.
Ω_X	Soluryhmän X mahdollisten kokonaistilojen joukko.
$\Pr[Q]$	Tapahtuman Q todennäköisyys.
$\Pr[Q\|R]$	Tapahtuman Q todennäköisyys ehdolla R.
$p_x(a) = p_x(a, t)$	Solun x todennäköisyys $\Pr[\sigma_x(t) = a]$ olla tilassa a hetkellä t.
$p_X(A) = p_X(A, t)$	Soluryhmän X todennäköisyys $\Pr[\sigma_X(t) = A]$ olla kokonaistilassa A hetkellä t.
$r_x(a \to b\|C)$	Tapahtumatiheys, jolla solu x siirtyy tilasta a tilaan b sen naapuruston N_x ollessa tilassa C. (Alaindeksi voidaan jättää pois, jollei sekaannuksen vaaraa ole.)
$r_X(A \to B\|C)$	Tapahtumatiheys, jolla soluryhmä X siirtyy tilasta A tilaan B sen naapuruston N_X ollessa tilassa C. (Alaindeksi voidaan jättää pois, jollei sekaannuksen vaaraa ole.)

Viitteet

[Ada02] Andrew Adamatzky, *Collision-based computing*, Springer, Berlin, 2002.

[Gar70] Martin Gardner, *Mathematical Games: The fantastic combinations of John Conway's new solitaire game "Life"*, Scientific American **223** (1970), 120–123.

[vB00] Minus van Baalen, *Pair approximations for different spatial geometries*, The Geometry of Ecological Interactions: Simplifying Spatial Complexity (Dieckmann, Law, and Metz, eds.), Cambridge University Press, 2000, pp. 359–387.

[Wol02] Stephen Wolfram, *A New Kind of Science*, Wolfram Media, 2002.